Balance of exploratory and reward-based behaviors in children

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in Psychology in the undergraduate colleges of The Ohio State University

By

Tim Horn

The Ohio State University

April 2020

Project Advisor: Dr. Vladimir Sloutsky and Nathan Blanco, Department of Psychology

Abstract

Making the best decision in a given situation requires a person to strike a unique balance between exploiting already known information and exploring new options. Previous research has shown that children have a greater tendency to explore than adults. Exploration is a long-term strategy that is utilized to gain more information about unknown options, sacrificing immediate rewards for information that can be used to achieve better outcomes later. This study pitted this long-term strategy versus immediate reward to see if children's unique balance gave them an advantage over adults in a situation where higher levels of exploration are necessary to succeed. To examine this possibility, a sample of 6 to 10-year-old children ($N = 78$) and adults ($N = 56$) performed 150 trials of a two-choice task with a complex reward structure in which immediate exploitation was disadvantageous. Surprisingly, children performed as well as adults in this task through all 150 trials. Separating the task into three 50-trial blocks revealed children's exploration tendency led them to outperform adults at the beginning of the task. Additionally, both children and adults performed better as the experiment progressed while children's exploration tendency decreased and adult's remained constant throughout the study. These results suggest the theory that adults have better decision-making skills is incomplete. Instead, children and adults have different default-modes of decision-making that gives children an advantage in a situation that requires exploration to be successful. Adults' advantage later in the experiment can be explained by children having an underdeveloped prefrontal cortex, which is necessary to learn and adapt an appropriate strategy after gaining the information necessary from exploration earlier in the experiment.

Balance of exploratory and reward-based behaviors in children

Exploration is a crucial activity for any organism to learn and make effective decisions (Hills et al., 2015; Melhorn et al., 2015). Determining to explore a new choice or to exploit a familiar option is a key component to daily decision-making (March, 1991). For example, when going to a favorite ice cream shop, a decision must be made between selecting a previously tried flavor with a known and high desirability (i.e. exploit) or a new ice cream flavor with unknown desirability (i.e. explore). An individual who exclusively exploits a previously tried flavor will fail to recognize the superiority of another option. The individual who exclusively explores new flavors will inevitably accumulate an opportunity cost by trying new flavors that may be less desirable than the highly desirable option. The most effective decision-making requires a unique balance of exploitation and exploration that varies depending on the situations and an individual's experience and goals (Daw et al., 2006; Blanco et al., 2015).

Young children are in a situation where they both know little and their direct needs are provided for by parents. If children know little about their environment, exploration would benefit them more than exploitation. This is because there is more potential knowledge acquisition through the exploration process (Sutton, 1998; Mata et al., 2013). There would be less need for children to exploit because of direct needs being provided for by parents. This situation suggests children should be shifted toward exploration to acquire knowledge of their surroundings. In contrast, adults already have knowledge and make rewarding decisions to attend to their needs. These two situations suggest that the balance of exploration and exploitation should change over the course of development.

Two research groups developed tasks to test the theory that children and adults have varying balances of exploration and exploitation. Schulz et al. (2019) had older children and adults make choices when there a high volume of options available. The goal of the task was to maximize the rewards received. One condition of the task organized masked and unmasked (i.e. visible value) options in spatially correlated areas which required an initial strategy of exploration to gain knowledge of the reward distribution. The study found that children explored far more than adults while underperforming exploitative adults. Blanco and Sloutsky (2019) conducted a similar four-choice task with four and five-year-old children and adults. The objective was to maximize rewards in a stable reward output structure. Similar to the Schulz et al. findings, Blanco and Sloutsky found that children showed heightened levels of exploration while adults had a reward-directed, exploitation tendency. This provides initial support for differential decision-making goals and strategies across development.

The optimal balance of exploration and exploitation also depends on the task at hand. An environment that is stable and well-understood needs less exploration. In contrast, situations where the environment is unknown or changing over time would require more exploration. For instance, an ice cream shop decides to rotate out all of their flavors every three months. Frequent customers of the ice cream shop should now explore flavors with unknown desirability to find new ice cream flavors due to the rotations.

The goal of the present study is to understand how the process of exploration develops in a person over time. Development is often assumed as the process of fine-tuning cognitive functioning that is both continuous and linear into adulthood. An alternative perspective is that children and adults simply use their cognitive processes differently due to variations in their goals and strategies as described previously. Certain situations geared more towards children's cognitive strengths may allow them to outperform adults. Specifically, it is possible that in certain circumstances an adult's exploitation tendency can lead them to underperform young children due to their balance shifted toward exploration.

The particular task used in the current experiment is adapted from Gureckis and Love's Farming on Mars task (2009). Versions of this task have been widely used for other aspects of decision-making (Cooper et al., 2013; Otto, Markman & Love, 2012; Pang, Otto & Worthy, 2014). This task was chosen because the reward output is altered after each choice made by the participant. Since the output of the reward structure changes over time, it is critical to explore, examine and learn from both options to perform well on the task. Very few adults performed well in the Farming on Mars task. Much of this poor performance can be attributed to their failure of exploring enough because of their tendency to exploit the immediately appealing option (Gureckis & Love, 2009). One account predicts young children's shift towards exploration may lead them to outperform exploitative adults since the task requires exploring the options to understand the function. A contrasting prediction is that children should be worse at the current task due to worse overall top-down control resulting from an underdeveloped prefrontal cortex (Bunge et al., 2002; Casey, Giedd, & Thomas, 2000; Sowell, et al. 1999). In this instance, the children would fail to learn anything from the additional exploration.

Method

Participants

Child participants from the greater Columbus, Ohio community were recruited from preschools and daycare centers based on returned parental permission forms. Child participants (*N* = 79) were six through ten-years-old (44 boys and 35 girls, *M* = 7.35 years) One child attrited out of the study after completing approximately 50 trials. Removing this participant from the data analysis did not impact any statistical analyses performed but was reported in efforts to report all available data from participants. All children received many stickers and a $10 Target gift card for participation. A total of 56 adults (27 men, 27 women and 2 chose not to respond, *M* = 19.72 years) from The Ohio State University participated in the study for partial course credit.

Materials

The experiment was run on a standard laptop computer using a custom code written in Python. A touchscreen was used with child participants. The laptop computer was placed approximately 50 cm away from the participant. All instructions and stimuli were displayed on the laptop. Figure 1 illustrates the stimuli used for this task. Participants were tested in a quiet room on campus in one session.

Design

The task was a simplified, child-friendly version of the Farming on Mars task named Candy Mining on Saturn (Gureckis & Love, 2009). What discriminates the task from the Farming on Mars task was its child-friendly framing, stimuli and reward payout ranges with familiar numbers for young children. The game was framed as taking place in a cave on an alien planet to mine for candy. All participants had two choices to choose from (Figure 1) to mine for candy. One mining tool choice was designed to be the better option in the long-term although it appeared worse in the short-term. Each tool gave back a predetermined amount of candy based on the amount of short and long-term choices in the last ten trials (Figure 2). The more long-term options the participant chose, the higher the reward was for both options. Specifically, this means the most optimal strategy is to always select the long-term option. The locations of the tool and the reward output calculation remained stable for the entire experiment. There was counterbalancing between participants of the short and long-term option appearing on either the left or right side of the screen as shown in Figure 1.

Each participant was informed their goal was to earn as much candy as possible on the game so it can later be exchanged for stickers. Participants choose one tool on each of the 150 trials. Children used a touchscreen to make each choice while adults did so with mouse presses. Each time the participant made a choice, a screen showing the amount of candy won appeared and stayed on the screen for 1.5 seconds. The total amount of candy earned was represented as a meter and was updated after each selection. Participant saw a screen that indicated they won a sticker after earning 250 candies.

Figure 1. Stimuli for two-choice alternative with the cave background used. The meter is a function of how many candies received in total and rises with each press in accordance with the

output number of candies. The lines are benchmarks for every 250 pieces of candy earned which translates to one sticker.

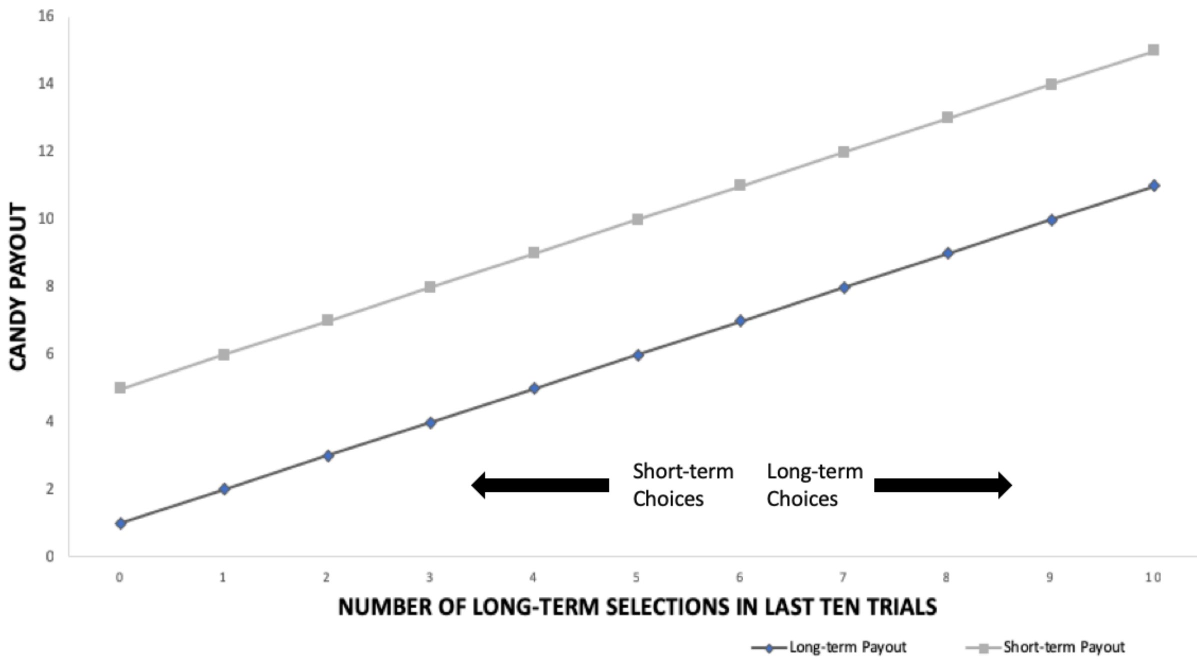

Figure 2. Payout function for the Candy Mining task adapted from Gureckis and Love for younger children (2009). The horizontal axis shows the amount of times the long-term tool was selected in the last ten trials. The vertical axis shows the amount of candy won based on which mining tool was selected. The candy payout for the short-term option is better than the long-term option on every trial. Each short-term selection moves the payout amount of both short and long-term selections one data point to the left, and a long-term selection moves the payout amount one to the right. The most optimal strategy is to always select the long-term option.

Results

Long-term Selections. The main measure of performance was the proportion of long-term option selections made by children and adults in each of the three 50-trial blocks (Figure 3). A 3 X 2 (50-Trial Block X Age Group) repeated-measures ANOVA found a main effect of block, $F(2,131) = 60.095$, $p < .001$, no main effect of age, $F(1,131) = .039$, $p = .844$ and found an interaction between block X age, $F(2,131) = 12.052$, $p = .001$. In order to better understand the interaction, analyses were done to determine the difference in the proportion of long-term responses of adults and children by individual block. Average correct long-term responses in Block 1 were higher for children ($M = 36.40\%$) than adults ($M = 27.50\%$), $t(133) = 2.874$, $p = .005$, 95% CI [$.028 \leq (\mu_1 - \mu_2) \leq .151$], $d = .507$. Block 2 had no significant differences between children ($M = 43.95\%$) and adults ($M = 43.30\%$), $t(132) = .136$, $p = .892$, 95% CI [$-.100 \leq (\mu_1 - \mu_2) \leq .087$], $d = .023$ and Block 3 showed no difference between adults ($M = 53.55\%$) and children ($M = 46.16\%$), $t(132) = 1.635$, $p = .104$, 95% CI [$-.016 \leq (\mu_1 - \mu_2) \leq .163$], $d = .286$. This along with the interaction between block X age shows that children outperformed adults in the beginning of the experiment. A 3 X 1 repeated-measures ANOVA revealed a main effect of block in children, $F(2,77) = 6.779$, $p = .002$ and in adults $F(2,55) = 32.163$, $p > .001$. This suggests that both children and adults improved as the experiment progressed.

Average Switch Frequency. Switch frequency was used to determine the levels of exploration seen in participants (Figure 4). In this experiment, switch frequency is defined as the proportion a participant selected a different option than they did on the previous trial. Switch frequency was calculated for children and adults in each of three 50-trial blocks. A 3 X 2 (50-Trial Block X Age Group) repeated-measures ANOVA found an effect of block, $F(2,131) = 6.094$, $p = .003$, a main effect of age, $F(2,131) = 7.513$, $p = .007$, and an interaction between block X age, $F(1,131) = 5.192$, $p = .024$. A 3 X 1 repeated-measures ANOVA revealed a main effect of block in children, $F(2,77) = 11.026$, $p > .001$ and no main effect of block in adults, $F(2,55) = .583$, $p = .560$. The main effect of block shows a decrease in switch frequency over the course of the experiment for children, while there was no change in adults. Additionally, the main effect of age suggests that children have a higher switch frequency than adults. Further analyses were done to determine differences in switch frequency by adults and children in each individual block. Average switch frequency for Block 1 was higher for children (*M* = 45.10%) than adults (*M* = 28.26%), $t(133) = 4.488$, $p < .001$, 95% CI [$.094 \leq (\mu_1 - \mu_2) \leq .243$), $d = .82$. Block 2 also showed higher switch frequency in children (*M* = 34.87%) than adults (*M* = 25.57%), $t(132) = 2.075$, $p = .04$, 95% CI [$.004 \leq (\mu_1 - \mu_2) \leq .182$], $d = .371$. Block 3 showed no difference between children (*M* = 34.17%) and adults (M = 28.29%), $t(132) = 1.183$, $p = .239$, 95% CI [$-.040 \leq (\mu_1 - \mu_2) \leq .157$], $d = .211$. Overall, children had a higher switch frequency than adults throughout the experiment. Children's switch frequency decreased as the experiment progressed while adults switch frequency remained unchanged.

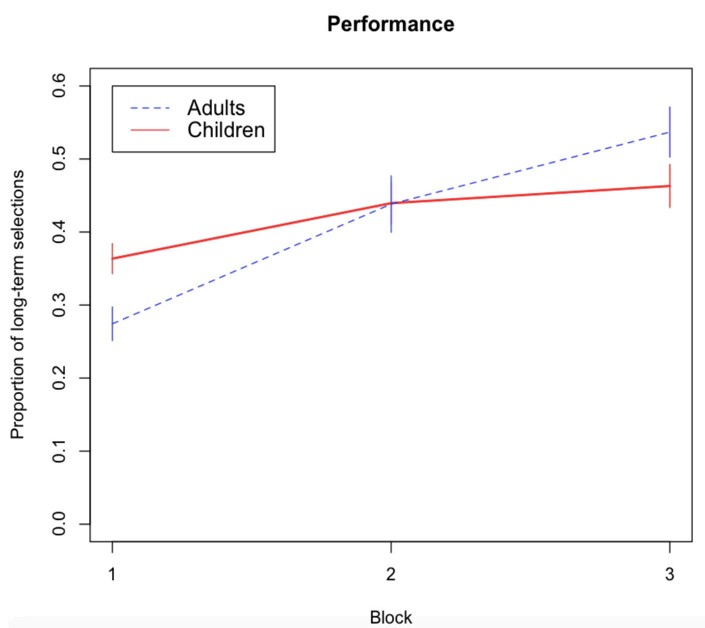

Figure 3. Performance on task was operationalized as the proportion of long-term selections which is found on the Y-axis. Additionally, one block = 50 trials. Children outperformed adults at the beginning of the experiment, while adults appear to be doing numerically better than children at the end. Error bars represent standard errors.

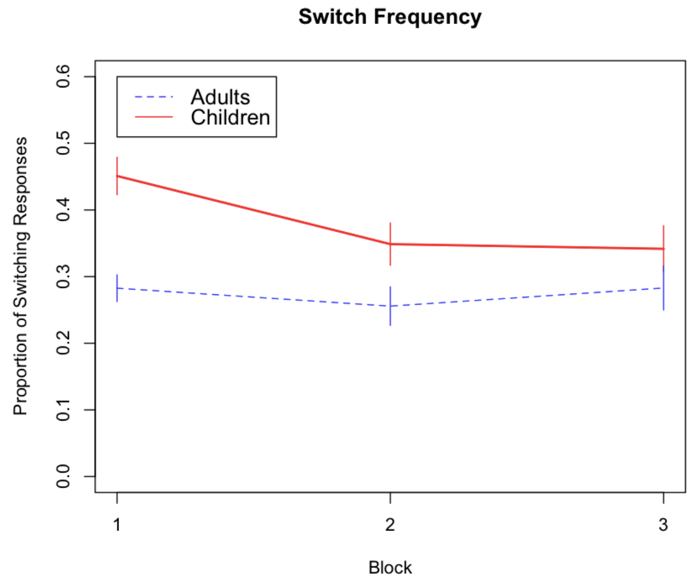

Figure 4. Switch frequency proportion. A switch was defined in this task as selecting a different option than the one selected in the previous trial. Additionally, one block = 50 trials. Children had a higher switch frequency than adults throughout the experiment. Children's switch frequency decreased by block while adult's remained constant throughout. Error bars represent standard errors.

Discussion

People are faced with decisions that often require them to lean towards either exploiting an already known option or exploring an unknown option, and the situation determines which strategy would be the most beneficial (Daw et al., 2006; Blanco et al., 2015). The present study served two purposes. First, it was done to test and extend prior findings that suggested children's default-mode of decision-making was shifted towards exploration and adults were shifted towards exploitation (Blanco & Sloutsky, 2019). Second, this study challenged the idea that decision-making processes are simply refined by development and should be at their most effective levels once adulthood is reached. An alternative hypothesis is that children and adults have different decision-making strategies due to their needs and circumstances.

In the present study, children and adults were given a two-choice task that required a high level of exploration to understand the complex reward output in which exploitation would lead to worse performance. To be successful in the Candy Mining task, it was key to explore both options adequately before deciding which option is the best to exploit. In this experiment, children and adults had comparable levels of overall performance on the task. It was also found that children explored between the two options more than adults throughout the entirety of the experiment. Despite no differences in overall performance, children outperformed adults in the beginning 50 trials of the experiment and adults numerically outperformed children at the end of the experiment.

The low performance of adults at the beginning of the experiment coupled with low levels of exploration (i.e. switch frequency) suggests that their first employed strategy was to exploit the immediately rewarding short-term option. This was not the case for children as the strategy they used was to explore the options more frequently. Due to these varying default decision-making strategies, children were able to outperform adults during the beginning stages of this novel task. As the experiment progressed, children's levels of exploration decreased and approximated adult levels by the end of the experiment. This trend suggests a potential strategy formation in children that is biased toward *directed* exploration (i.e. using knowledge of the environment to plan exploration) instead of random or undirected exploration (Blanco et al., 2013; Wilson et al., 2014; Blanco et al., 2015). While this task cannot definitively confirm the existence of directed exploration, it raises the possibility that this process emerges well before previous research suspected. For example, Somerville et al. (2017) documented directed exploration in adolescence, but it is rare to see any evidence of it earlier in development.

Unfortunately, there are limitations to a child's ability to learn from the new information they gained at the beginning of the experiment. This may be due to an underdeveloped prefrontal cortex (Bunge et al., 2002; Casey, Giedd, & Thomas, 2000; Sowell, et al. 1999). Less development in the prefrontal cortex is associated with poorer ability in processes such as short-term memory that is critical for learning from additional exploration. More specifically to the study, it is possible that the reward structure of the Candy Mining task was too complex for children to utilize the new information gained from exploration on subsequent trials.

The results of this study, particularly within the first block, indicate that the theory of decision-making processes becoming more efficient with age should be revised. A possible alternative theory would suggest that children and adults simply have varying default balances of exploratory and reward-seeking behaviors that best meet their unique circumstances in life. Specifically, children are more shifted towards exploration because they face a lot of unknowns in the world and do not have to worry about the cost, and potential risks, of exploration due to having their basic needs provided for by adults. On the other hand, adults are in a situation where they can use knowledge acquired to obtain maximum rewards. Naturally, exploratory children will discover desirable options that are worth exploiting over their development that will start to shift their tendency towards exploitation.

This study shows that in the early phases of a changing situation, children's long-term strategy of exploration may lead to better outcomes than an adult's exploitation for immediate rewards. Future iterations of the Candy Mining task should consider adding more than 150 trials to assess if more time with the task facilitates more learning. Additionally, adjusting the complexity of the reward structure to make the task easier would test the robustness of the current findings. New tasks and paradigms should work to uncover if there is a critical time period in development where the tendency of exploration shifts toward a tendency to exploit.

CONCLUSION

Children and adults are in different life circumstances that require differing default balances of exploration and exploitation in their decision-making to meet their needs. Specifically, this means that most children have a greater tendency towards exploration than adults. It is oversimplified to view children's decision-making skills inferior to adults' because a child's goal and current circumstance in life are much different from those of an adult. This study provides some of the first evidence that children's decision-making strategies can be better than adults' in the initial stages of a changing situation.

References

Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G., & Love, B. C. (2013). The influence of depression symptoms on exploratory decision-making. Cognition, 129, 563–568.

Blanco, N. J., Love, B. C., Cooper, J. A., McGeary, J. E., Knopik, V. S., & Maddox, W. T. (2015). A frontal dopamine system for reflective exploratory behavior. *Neurobiology of Learning and Memory, 123*, 84-91.

Blanco, N. J., & Sloutsky, V. (2019, August 5). Systematic exploration and uncertainty dominate young children's choices. *PsyArXiv*. https://doi.org/10.31234/osf.io/72sfx

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, 33(2), 301-311.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, 54(1-3), 241-257

Cooper, J. A., Worthy, D. A., Gorlick, M. A., & Maddox, W. T. (2013). Scaffolding across the lifespan in history-dependent decision-making. *Psychology and Aging,28*(2), 505-514. doi:10.1037/a0032717

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876.

Gureckis, T. M., & Love, B. C. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition, 113*(3), 293-313. doi:10.1016/j.cognition.2009.03.013

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences, 19*(1), 46-54.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science, 2*(1), 71-87. doi:10.1287/orsc.2.1.71v

Mata R, Wilke A, Czienskowski U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Front. Neurosci.* 7:53. https://doi.org/10.3389/fnins.2013.00053

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., & Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision, 2*(3), 191.

Otto, A. R., Markman, A. B., & Love, B. C. (2011). Taking More, Now. *Social Psychological and Personality Science, 3*(2), 131-138. doi:10.1177/1948550611411311

Pang, B., Otto, A. R., & Worthy, D. A. (2014). Self-Control Moderates Decision-Making Behavior When Minimizing Losses versus Maximizing Gains. *Journal of Behavioral Decision Making, 28*(2), 176-187. doi:10.1002/bdm.1830

Schulz, E., Wu., C. M., Ruggeri, A., & Meder B. (2019). Searching for rewards like a child means less generalization and more directed exploration**.** *Psychol Sci,* 30(11), 1561-1572. doi: 10.1177/0956797619863663.

Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Akar, N. A., Insel, C., & Wilson, R. C. (2017*).* Charting the expansion of strategic exploratory behavior during adolescence. Journal of Experimental Psychology*: General,* 146(2), 155–164. doi: 10.1037/xge0000250

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, 2(10), 859-861.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. doi: 10.1037/a0038199